Lee

ENTERED

```
           SEQUENCE LISTING
  4 (1) GENERAL INFORMATION:
  6      (i) APPLICANT: Boyd, Michael R.
  7                     Gustafson, Kirk R.
  8                     Shoemaker, Robert H.
  9                     McMahon, James B.
 11     (ii) TITLE OF INVENTION: ANTIVIRAL PROTEINS AND PEPTIDES, DNA
 12                     CODING SEQUENCES THEREFOR, AND USES THEREOF
 14    (iii) NUMBER OF SEQUENCES: 4
 16     (iv) CORRESPONDENCE ADDRESS:
 17          (A) ADDRESSEE: Leydig, Voit & Mayer, Ltd.
 18          (B) STREET: Two Prudential Plaza, Suite 4900
 19          (C) CITY: Chicago
 20          (D) STATE: IL
 21          (E) COUNTRY: U.S.A.
 22          (F) ZIP: 60601-6780
 24      (v) COMPUTER READABLE FORM:
 25          (A) MEDIUM TYPE: Floppy disk
 26          (B) COMPUTER: IBM PC compatible
 27          (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 28          (D) SOFTWARE: PatentIn Release #1.0, Version #1.25
 30     (vi) CURRENT APPLICATION DATA:
 31          (A) APPLICATION NUMBER: US/09/427,873
 32          (B) FILING DATE: 27-Oct-1999
 33          (C) CLASSIFICATION:
 35    (vii) PRIOR APPLICATION DATA:
 36          (A) APPLICATION NUMBER: 08/429,965
 37          (B) FILING DATE:
 39   (viii) ATTORNEY/AGENT INFORMATION:
 40          (A) NAME: Larcher, Carol
 41          (B) REGISTRATION NUMBER: 35243
 42          (C) REFERENCE/DOCKET NUMBER: 61037
 44     (ix) TELECOMMUNICATION INFORMATION:
 45          (A) TELEPHONE: (312)616-5600
 46          (B) TELEFAX: (312)616-5700
 49 (2) INFORMATION FOR SEQ ID NO: 1:
 51      (i) SEQUENCE CHARACTERISTICS:
 52          (A) LENGTH: 327 base pairs
 53          (B) TYPE: nucleic acid
 54          (C) STRANDEDNESS: double
 55          (D) TOPOLOGY: linear
 57     (ii) MOLECULE TYPE: DNA (genomic)
 60     (ix) FEATURE:
 61          (A) NAME/KEY: CDS
 62          (B) LOCATION: 10..312
 65     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
 67 CGATCGAAG CTT GGT AAA TTC TCC CAG ACC TGC TAC AAC TCC GCT ATC          48
```

C--> 31
C--> 32

```
         Input Set : A:\09427873.txt
         Output Set: N:\CRF3\08012000\I427873.raw

68               Leu Gly Lys Phe Ser Gln Thr Cys Tyr Asn Ser Ala Ile
69                 1               5                  10
71 CAG GGT TCC GTT CTG ACC TCC ACC TGC GAA CGT ACC AAC GGT GGT TAC      96
72 Gln Gly Ser Val Leu Thr Ser Thr Cys Glu Arg Thr Asn Gly Gly Tyr
73      15                  20                  25
75 AAC ACC TCC TCC ATC GAC CTG AAC TCC GTT ATC GAA AAC GTT GAC GGT      144
76 Asn Thr Ser Ser Ile Asp Leu Asn Ser Val Ile Glu Asn Val Asp Gly
77   30              35                  40                  45
79 TCC CTG AAA TGG CAG CCG TCC AAC TTC ATC GAA ACC TGC CGT AAC ACC      192
80 Ser Leu Lys Trp Gln Pro Ser Asn Phe Ile Glu Thr Cys Arg Asn Thr
81          50                  55                  60
83 CAG CTG GCT GGT TCC TCC GAA CTG GCT GCT GAA TGC AAA ACC CGT GCT      240
84 Gln Leu Ala Gly Ser Ser Glu Leu Ala Ala Glu Cys Lys Thr Arg Ala
85          65                  70                  75
87 CAG CAG TTC GTT TCC ACC AAA ATC AAC CTG GAC GAC CAC ATC GCT AAC      288
88 Gln Gln Phe Val Ser Thr Lys Ile Asn Leu Asp Asp His Ile Ala Asn
89          80                  85                  90
91 ATC GAC GGT ACC CTG AAA TAC GAA TAACTCGAGA TCGTA                     327
92 Ile Asp Gly Thr Leu Lys Tyr Glu
93      95                  100
96 (2) INFORMATION FOR SEQ ID NO: 2:
98      (i) SEQUENCE CHARACTERISTICS:
99          (A) LENGTH: 101 amino acids
100         (B) TYPE: amino acid
101         (D) TOPOLOGY: linear
103     (ii) MOLECULE TYPE: protein
105     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
107 Leu Gly Lys Phe Ser Gln Thr Cys Tyr Asn Ser Ala Ile Gln Gly Ser
108   1               5                  10                  15
110 Val Leu Thr Ser Thr Cys Glu Arg Thr Asn Gly Gly Tyr Asn Thr Ser
111             20                  25                  30
113 Ser Ile Asp Leu Asn Ser Val Ile Glu Asn Val Asp Gly Ser Leu Lys
114             35                  40                  45
116 Trp Gln Pro Ser Asn Phe Ile Glu Thr Cys Arg Asn Thr Gln Leu Ala
117         50                  55                  60
119 Gly Ser Ser Glu Leu Ala Ala Glu Cys Lys Thr Arg Ala Gln Gln Phe
120   65                  70                  75                  80
122 Val Ser Thr Lys Ile Asn Leu Asp Asp His Ile Ala Asn Ile Asp Gly
123             85                  90                  95
125 Thr Leu Lys Tyr Glu
126             100
128 (2) INFORMATION FOR SEQ ID NO: 3:
130     (i) SEQUENCE CHARACTERISTICS:
131         (A) LENGTH: 327 base pairs
132         (B) TYPE: nucleic acid
133         (C) STRANDEDNESS: double
134         (D) TOPOLOGY: linear
136     (ii) MOLECULE TYPE: DNA (genomic)
139     (ix) FEATURE:
```

RAW SEQUENCE LISTING   DATE: 08/01/2000
PATENT APPLICATION: US/09/427,873   TIME: 15:07:56

Input Set : A:\09427873.txt
Output Set: N:\CRF3\08012000\I427873.raw

```
140          (A) NAME/KEY: CDS
141          (B) LOCATION: 1..327
144     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
146 GAC TAC AAG GAC GAC GAT GAC AAG CTT GGT AAA TTC TCC CAG ACC TGC        48
147 Asp Tyr Lys Asp Asp Asp Asp Lys Leu Gly Lys Phe Ser Gln Thr Cys
148  1               5                  10                  15
150 TAC AAC TCC GCT ATC CAG GGT TCC GTT CTG ACC TCC ACC TGC GAA CGT        96
151 Tyr Asn Ser Ala Ile Gln Gly Ser Val Leu Thr Ser Thr Cys Glu Arg
152          20                  25                  30
154 ACC AAC GGT GGT TAC AAC ACC TCC TCC ATC GAC CTG AAC TCC GTT ATC       144
155 Thr Asn Gly Gly Tyr Asn Thr Ser Ser Ile Asp Leu Asn Ser Val Ile
156          35                  40                  45
158 GAA AAC GTT GAC GGT TCC CTG AAA TGG CAG CCG TCC AAC TTC ATC GAA       192
159 Glu Asn Val Asp Gly Ser Leu Lys Trp Gln Pro Ser Asn Phe Ile Glu
160          50                  55                  60
162 ACC TGC CGT AAC ACC CAG CTG GCT GGT TCC TCC GAA CTG GCT GCT GAA       240
163 Thr Cys Arg Asn Thr Gln Leu Ala Gly Ser Ser Glu Leu Ala Ala Glu
164  65                  70                  75                  80
166 TGC AAA ACC CGT GCT CAG CAG TTC GTT TCC ACC AAA ATC AAC CTG GAC       288
167 Cys Lys Thr Arg Ala Gln Gln Phe Val Ser Thr Lys Ile Asn Leu Asp
168          85                  90                  95
170 GAC CAC ATC GCT AAC ATC GAC GGT ACC CTG AAA TAC GAA               327
171 Asp His Ile Ala Asn Ile Asp Gly Thr Leu Lys Tyr Glu
172          100                 105
175 (2) INFORMATION FOR SEQ ID NO: 4:
177     (i) SEQUENCE CHARACTERISTICS:
178          (A) LENGTH: 109 amino acids
179          (B) TYPE: amino acid
180          (D) TOPOLOGY: linear
182     (ii) MOLECULE TYPE: protein
184     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:
186 Asp Tyr Lys Asp Asp Asp Asp Lys Leu Gly Lys Phe Ser Gln Thr Cys
187  1               5                  10                  15
189 Tyr Asn Ser Ala Ile Gln Gly Ser Val Leu Thr Ser Thr Cys Glu Arg
190          20                  25                  30
192 Thr Asn Gly Gly Tyr Asn Thr Ser Ser Ile Asp Leu Asn Ser Val Ile
193          35                  40                  45
195 Glu Asn Val Asp Gly Ser Leu Lys Trp Gln Pro Ser Asn Phe Ile Glu
196          50                  55                  60
198 Thr Cys Arg Asn Thr Gln Leu Ala Gly Ser Ser Glu Leu Ala Ala Glu
199  65                  70                  75                  80
201 Cys Lys Thr Arg Ala Gln Gln Phe Val Ser Thr Lys Ile Asn Leu Asp
202          85                  90                  95
204 Asp His Ile Ala Asn Ile Asp Gly Thr Leu Lys Tyr Glu
205          100                 105
```

VERIFICATION SUMMARY   DATE: 08/01/2000
PATENT APPLICATION: US/09/427,873   TIME: 15:07:57

Input Set : A:\09427873.txt
Output Set: N:\CRF3\08012000\I427873.raw

L:31 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
L:32 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]